Evaluating News Media Reports On the 'Blue Whale Challenge' For Adherence to Suicide Prevention Safe Messaging Guidelines

REBECCA ROTH, Clemson University, USA
JACLYN ABRAHAM, University of Central Florida, USA
HEIDI ZINZOW, Clemson University, USA
PAMELA WISNIEWSKI, University of Central Florida, USA
AMRO KHASAWNEH, Clemson University, Clemson, USA
KAPIL CHALIL MADATHIL, Clemson University, USA

Corresponding Author: Rebecca Roth, raroth@g.clemson.edu

Warning: This paper discusses suicide and self-harming behaviors on social media. Therefore, the reader should use their discretion as to whether they want to proceed.

The Blue Whale Challenge (BWC) is an online viral "game" that allegedly encourages youth and young adults towards self-harming behaviors that could eventually lead to suicide. The BWC can be situated within a larger phenomenon of viral online self-harm challenges, which may be propagated through both social media and news sources. Research has established that suicide is a global public health issue that is known to be influenced by media reporting. Violation of safe messaging guidelines has been shown to increase imitative suicides, particularly in youth and young adults. Given the confirmed effects of news media reporting, we analyzed 150 digital newspaper articles reporting on the BWC to assess whether they adhered to suicide prevention safe messaging guidelines. Overall, 81% of the articles violated at least one contagion-related guideline, most commonly normalizing suicide, discussing means of suicide, and sensationalizing. Even though the majority (91%) of the articles adhered to at least one health-promotion guideline, such as emphasizing prevention, the articles did not follow these guidelines on a deep and comprehensive level. Through thematic analysis, we also found evidence of potential misinformation in reporting, where the articles unequivocally attributed many suicides to the BWC with little or no evidence. Additionally, articles often stated an individual's reason for participating in the challenge without interviewing the individual or those close to the individual, another aspect of potential misinformation due to lack of evidence. A contribution of the current study is the synthesis of safe messaging guidelines that can be used in future research. This study contributes to the understanding of news reporting practices regarding suicide and self-harm in regard to the BWC and similar online challenges. We discuss how sensationalized news media reports on the BWC could unintentionally propagate suicide contagion effects that normalize self-harming behaviors among youth. We then examine implications for practice and policy, such using automated approaches to aid reporters in adhering to safe messaging guidelines.

Authors' addresses: REBECCA ROTH, Clemson University, Clemson, SC, 29634, USA, raroth@g.clemson.edu,; JACLYN ABRAHAM, University of Central Florida, Orlando, FL, 32816, USA, Jaclyn.abraham@knights.ucf.edu; HEIDI ZINZOW, Clemson University, Clemson, SC, 29634, USA, hzinzow@clemson.edu; PAMELA WISNIEWSKI, University of Central Florida, Orlando, FL, 32816, USA, pamela.wisniewski@ucf.edu; AMRO KHASAWNEH, Clemson University, Clemson, SC, 29634, USA, , akhasaw@g.clemson.edu; KAPIL CHALIL MADATHIL, Clemson University, Clemson, SC, 29634, USA, kmadath@clemson.edu

CCS CONCEPTS

• Collaborative and social computing • Collaborative and social computing theory, concepts, paradigms • Social content sharing

KEYWORDS

Suicide, Non-Suicidal Self-Injury, Social Media Challenges, Blue Whale Challenge

1 INTRODUCTION

As of 2016, suicide is the second leading cause of death among teens and young adults in the United States [60,86]. Some research suggest that increased screen time and social media usage is a factor related to the increase in depressive symptoms and suicide rates [85], but other researchers have argued that this correlative relationship does not imply causation [77] and/or is overstated [66]. Indeed, social media usage is on the rise within young adults and adolescent populations [8]. Ninety-Five percent of teenagers report that they have access to a smartphone and use a larger number of platforms than in previous years. Meanwhile, Pew Research reports that youth and young adults have mixed feelings about the impact of social media on their lives [8]. Teenagers claim that while there are benefits such as connecting with friends and family, there are negative effects such as rumor spreading, hurting relationships, and providing an unrealistic view of others' lives [8]. Given both the potential benefits and harm social media poses for our youth, further research is warranted to explore various aspects of social media use that may tip this delicate balance in unanticipated ways. Embedded in this context of youth suicide, well-being, and social media use, we studied a recent phenomenon called the Blue Whale Challenge (BWC).

In the BWC, like other viral social media challenges, participants record or take pictures of themselves engaging in specific activities and share their experience through social media [79]. The BWC falls in a subset of these challenges that involve self-harming behavior. Given the viral nature of social media, challenges such as the BWC pose potential physical and psychological risks for those who participate [42]. Mukhra et al. [57] described the BWC as a "deadly online craze," (p.6) that involves engaging the participant in a series of 50 tasks, which escalate from self-harming behaviors (e.g., putting oneself into risky situations such as standing on the edge of a tall building, cutting oneself) to eventual suicide. In 2019, Lupariello et al. [54] conducted a case study on five suspected cases of the BWC in a pediatric unit that handles child abuse cases. Their study was one of the first to document actual evidence that the BWC is fact, rather than fiction (e.g., pictures of youth self-harm). While some question whether or not suicides have been attributed to the challenge [96], numerous reports have surfaced across both social and news media regarding the BWC.

Since news reporting on suicide has been shown to influence imitative suicides [39,41,70], researchers have developed safe messaging guidelines [17,35,62] in order to prevent suicide contagion effects. Because the BWC involves suicidal behavior, it is possible that news media reporting on the BWC might also influence the propagation of online self-harm challenge participation in a similar fashion. Therefore, we analyze 150 digital news media reports about the BWC to assess whether these reports adhered to safe messaging guidelines for suicide prevention. We contribute to the human computer interaction literature on social media and self-harm by applying suicide prevention safe messaging guidelines to a new area of concern, digital news

reports of online challenges that involve self-harm and/or suicide. The over-arching research questions below guided our study:

> **RQ1:** How does news media reporting on the BWC adhere to or violate contagion-related safe messaging guidelines?
> **RQ2:** How does news media reporting on the BWC adhere to or violate health promotion guidelines?
> **RQ3:** What are some problematic emergent themes regarding BWC news media reporting that can inform future safe messaging guidelines, research, and practice?

To answer these research questions, we first synthesized evidence-based safe messaging guidelines related to suicide prevention to create a conceptually driven framework divided into contagion-related and health promotion safe messaging guidelines. Next, we used an iterative approach to qualitatively analyze 150 news articles related to the BWC based on this framework. Specifically, we evaluated these news articles based on 12 contagion-related suicide prevention guidelines and six health promotion guidelines to assess whether and how they adhered to or violated these guidelines. Finally, we conducted a grounded thematic analysis of the news articles to identify emergent themes not included in our initial analysis that uncovered additional problematic news reporting practices that could potentially lead to unintended suicide contagion effects.

Overall, we found that most articles (81%) violated one or more contagion related guidelines. Many articles normalized suicide, gave specific means by which an individual attempted suicide, and sensationalized the death. Additionally, the majority of articles adhered at least one health promotion guideline (91%), and most articles emphasized prevention. Although many articles had a positive aspect to reporting, contagion-based reporting overshadowed health promotion reporting. Through our thematic analysis, we also identified two emergent themes related to online misinformation. First, articles often attributed deaths to the BWC with little or no evidence that confirmed causation. Second, many news articles made claims about the reason(s) that adolescents participated in the BWC without interviewing the individual or family. We discuss how the lack of confirmatory evidence being disseminated about the BWC may increase the potential for misinformation related to self-harm and suicide, which in turn, may increase the potential harm from contagion-related effects. The unique contributions of this research include: 1) A synthesized framework of conceptually driven and evidence-based safe messaging guidelines for suicide prevention that can be leveraged as a guide for reporting on suicide. This framework could be used by news reporters as well as researchers when reporting on self-harm, suicide, and online challenges and platforms that entail self-harming behaviors. 2) An empirical analysis assessing whether and how news reporting on the BWC adhered to or violated this framework of safe messaging guidelines. 3) Recommendations on how to improve suicide prevention safe messaging guidelines, journalist adherence, and the emergence of viral social media self-harm and suicide challenges, as well as implications for research and practice. The issue of news reporting on suicide is crucial due to the risk for imitative suicides. We provide both recommendations for news reporters, policy, and possible technological interventions in order to reduce potential contagion in the context of previous research. Our study both builds upon prior literature and extends it by applying established safe messaging guidelines for suicide prevention to the novel context of the BWC, a potentially harmful and viral social media suicide challenge.

2 Related Work

In the subsections below, we will first introduce behavioral contagion as the guiding theory behind our work. Next, we discuss the Blue Whale Challenge, contagion and the media, and suicide prevention safe messaging guidelines.

2.1 The Blue Whale Challenge

The Blue Whale Challenge (BWC) is an online challenge that started in Russia and spread throughout many countries, including India [57]. Mukhra et al. [57] describe the BWC in detail, including the 50 steps of the challenge, which start by isolating the victim, telling the individual to self-harm, preparing the individual to take their own life through increasingly risky tasks, and then asking the participant to attempt suicide. Additionally, the curators of the challenge typically ask for photographic evidence of these tasks online to make sure they have appropriate control over the victim [58].

Mukhra et al [58] describes the unique online nature of the challenge, where an administrator interacts with the participant solely via the internet. They frame the BWC as a public health crisis and a potential crime that has severe effects on teens. Yet, the validity of the BWC is equivocal as it is difficult to confirm deaths resulting from the challenge [21]. There is little academic and reliable information as to the origin, progression, or participation based on geographic location of the challenge. Khattar et al. [47] analyzed public posts and hashtags on VK, Instagram, and Twitter to identify patterns in how posts about the BWC spread through these networks. They found that quantifying the number of deaths actually associated with the BWC was "extremely difficult." Meanwhile, Lupariello et al. [54] conducted an in-depth case study on five suspected cases of the BWC in a pediatric unit that handles child abuse cases, including visual depictions of self-harm that appeared to result from the challenge. A recent study examined the number and valence of social media posts about the BWC and found that nearly a quarter of the posts advocated participation in the challenge [83]. As a result, the authors urged further research be conducted on the BWC. Regardless of whether deaths related to BWC can be confirmed, problematic media reporting on self-harm challenges such as BWC could result in imitative behavior and unintended contagion effects.

The published research on the BWC highlights the controversy regarding the validity of the BWC, its alleged association with youth suicides, and more importantly, raises concerns about how self-harm and suicide related to viral social media challenges should be reported and depicted in both the media and academic research. For instance, these articles included highly graphic images depicting self-harm [54] and outlined the exact steps needed to carry out the BWC [57]. While evidence-based research on the BWC is needed, one must also consider the ways in which we study this challenge and the possible negative effects of what we publish.

2.2 Self-Harm and Suicide Within Human-Computer Interaction Research

We situate our research on the BWC within the larger body of Computer Supported Collaborative Work (CSCW) and Human Computer Interaction research on social media, self-harm [37,64], and online safety [44,89–91]. The BWC can be classified as a type of digital self-harm [69] since it is an online phenomenon that encourages or promotes self-harming behavior that may also lead to suicide. Pater and Mynatt's foundational work on digital self-harm [69] describes both direct (e.g. cutting) and indirect self-harm (e.g. eating disorder behaviors) and highlights how those who view such behaviors could indirectly be promoting others to engage in the same or similar self-harming behaviors. The high comorbidity of eating disorders and self-harm show the importance of studying these behaviors in conjunction [30]. Pro-eating disorder communities have been linked to increased mental illness, highlighting the potentially harmful nature of these online communities [26]. Pater also makes the point that the online dissemination of harmful behaviors is more dangerous than the previous in person dissemination. Due to the size of the audience and

availability of posts across time, digital self-harm may have more detrimental effects than the offline version of the same behavior [68].

However, these communities can also be used as a source of social support for those struggling with mental illness; for example, Andalibi, Ozturk, and Forte [6] found that moderated self-harm oriented communities in which users could discuss their experiences anonymously and with less fear of stigma exhibited positive outcomes, such as members receiving needed validation and support from other members to cope with their mental illness. Similar, Choudhury and De [31] studied how Reddit users were able to express their negative emotions related to their experienced with mental illness and received social support from other users. Yet, Choudhury's [30] study on eating disorder communities on Tumblr found that there were more pro-eating disorder posts than pro-recovery posts. As such, Pater et al. [67] points out the need for social media platforms to redirect consumers and creators of self-harming content to resources that connect them to more positive communities, such as those that promote recovery.

Other human-computer interaction researchers have used machine learning techniques to better understand the nature of self-harm and suicide related posts on social media. For instance, Wang et al. [87] built a classification algorithm to detect self-harming behaviors within Flickr photos. They found that self-harming behaviors are more prevalent at night and that potential self-harm users are more active than normal social media users and receive a higher number of responses, even though they have fewer friends. Choudhury et al [32] used Reddit posts to predict when individuals discussing mental health would experience suicidal ideation. They found that factors such as increase in self-focus, lower semantic quality, lower engagement in the community, hopelessness, impulsivity, and loneliness predicted the shift to suicidal ideation among users. Researchers also found that banning search terms related to pro-eating disorder communities was ineffective because users simply changed their language slightly so that their content was no longer censored [27,68]. Therefore, these researchers pointed to the need to find better ways to reduce the spread of harmful content, such as potentially moderating these communities. The issue of censorship within online self-harm communities is controversial [25], and human computer interaction researchers continue to urge the CSCW community to further explore practices and policies to reduce the potentially negative effects of user generated digital self-harm content. Consequently, we join this effort by examining digital self-harm associated with the BWC and how the challenge has been depicted within digital news media articles. A contrast between online self-harm communities and digital news media reporting on the BWC is that online communities are often more insular [69], attracting like-minded individuals, rather than mainstream audiences. Online news reporting makes information about the BWC, and consequently self-harm, accessible to broad audiences, including vulnerable populations such as youth and individuals who suffer from mental illness. Therefore, understanding digital news media reports on the BWC makes a unique and important contribution to the CSCW literature on self-harm and suicide.

2.3 Digital News Media and Suicide Contagion Effects

The theory of behavioral contagion is based on Bandura's social learning theory, which states that people both mirror observed behaviors and change their attitudes toward that behavior upon observation [15]. Contagion refers to a behavior's ability to disseminate throughout a cohort of people [39]. Suicide contagion is one type of contagion, and is thus, placed within the category of behavioral contagion. This means that suicide can propagate throughout a group of people as observing suicidal behavior in another person can influence suicidal behavior in another [39]. Suicide contagion has been a well-studied phenomenon since the 1960's; there is a saturation of evidence that supports contagion through news media reports. Phillips et al. [72] call the presentation of suicide in the media "a natural advertisement for suicide," and recommended changing how suicide is covered, presented, and the placement of the message in order to reduce

imitative suicides [71]. Niederkrotenthaler et al. [28] found that variations in contagion depended on the quality and type of news story. For example, sensationalizing, propagating suicide myths, and reporting on suicide as an epidemic were associated with an increased risk of suicide [63]. This finding highlights the importance of responsible reporting practices.

Overall, the literature supports the presence of suicide contagion effects [55,84]. Many studies have provided compelling support for the effect of reporting on future suicides [39,41,70]. A common method for studying contagion is to observe clusters of suicides after high levels of attention, often media coverage, are given to a suicide. As such, studies have reported a positive correlation between media coverage on suicide and increases in suicide rates within a geographic area. The quality of news media reporting can influence contagion [13,61,62]. Specifically, poor quality reporting may lead to imitative suicide, whereas high quality reporting reduces this occurrence. In addition to the manner in which suicides are reported, other factors have also been associated with higher suicide contagion effects. For instance, Stack [78] found that celebrity suicides were over fourteen times as likely to lead to imitative suicides. However, there is less research across different communication mediums. This highlights the need for more research on suicide reporting in the media, both print and online, in order to minimize possible contagion effects. Further, this begs the question as to how emergent forms of modern-day media present new mediums and challenges for the viral spread of suicide contagion effects, especially as news media has now collided with user-generated content and social media. It is possible that problematic news reporting on viral self-harm challenges may result in contagion of self-harm challenge participation in a manner similar to suicide contagion. In this paper, we specifically focus on how news media reports discussing the BWC may contribute to this contagion effect based on whether and how they violate suicide prevention safe messaging guidelines. Next, we introduce and synthesize established safe messaging guidelines for suicide prevention.

3 Synthesizing Safe Messaging Guidelines for Suicide Prevention

Suicide prevention safe messaging guidelines have been developed as a response to the potential behavioral contagion of self-harm and suicide. Safe messaging guidelines were developed with the goal of reducing "copycat" suicides [82,94]. The need for suicide prevention safe messaging guidelines are empirically supported; for instance, a number of studies have found a decrease in imitative suicides over time after the implementation of such guidelines [17,35,62]. Safe messaging guidelines have been established cross-culturally across the world, including India [13,28,43], which is one of the leading countries reporting on the BWC. Prior research confirming suicide contagion effects of news media reporting [39,41,70] necessitates and has informed the development of safe messaging guidelines.

In this paper, we synthesize different, yet overlapping, safe messaging guidelines that have been widely accepted in practice and academic research (**Table 1**). To do this, one of our co-authors, an expert in suicide prevention research, first pointed us to established and evidence-based safe messaging guidelines for suicide prevention. We also performed a keyword search for other "suicide prevention safe messaging guidelines" to make sure our search was comprehensive. We chose these three sets of guidelines based on the authority of their authors and empirical evidence. These included the "Suicide Prevention Resource Center's (SPRC) Safe and Effective Messaging Guidelines" from the American Foundation for Suicide Prevention, "Recommendations for Reporting on Suicide" made by the Center for Disease Control, and the World Health Organization's "Preventing Suicide: A Resource for Media Professionals" [92].

The Suicide Prevention Resource Center's (SPRC) Safe and Effective Messaging Guidelines" were supported by the American Foundation for Suicide Prevention and partially written by one of the leading experts in suicide prevention, Dr. Madelyn Gould [82]. The "Recommendations for Reporting on Suicide" were established by the Center for Disease Control, American Association of Suicidality, American Foundation for Suicide Prevention, and many experts in the field. These guidelines are empirically based and the authors note they are based on over 50 worldwide studies on suicide contagion [94]. Lastly, The World Health Organization's "Preventing Suicide: A

Resource for Media Professionals" [92] were developed with the International Association for Suicide Prevention and a culturally diverse set of suicide prevention experts. These guidelines were also empirically based, citing literature on their prevention ability [92]. We omitted a fourth set of guidelines "Tips for Messaging Safely" [59] because they had not been empirically validated. Additionally, these guidelines had no information regarding experts authoring or endorsing them. This set of guidelines is distinct from the other three sets used because they are not specifically aimed at the factors that research shows curbs imitative suicide as compared with other empirically supported guidelines.

Next, we synthesized the three sets of guidelines by merging those which overlapped (e.g. emphasizing prevention [82,92,94]) and keeping those that were unique (e.g. interview suicide prevention experts [94]). In order to acknowledge that it is important to examine both health and disease, we divide guidelines into contagion related guidelines and health promotion guidelines. Just as the absence of disease is not health [4], the absence of harmful messaging is not positive messaging. Thus, we divide the types of messages into contagion-related guidelines (the negative to avoid) and health promotion guidelines (the positive to encourage). In **Table 1,** we include contagion-related guidelines on the left and health promotion guidelines to those which they most directly parallel on the right. The intersection of research on digital self-harm, suicide contagion effects, and safe-messaging guidelines collectively shows the importance of responsible digital news media reporting on the BWC. Thus, we evaluated whether news reporting on the BWC adheres to or violates safe messaging guidelines. Understanding patterns of adherence to safe messaging guidelines for the BWC will help inform future efforts to reduce contagion of online viral self-harm challenges.

3.1 Contagion-Related Guidelines

In **Table 1**, we present our synthesis of these guidelines grouped conceptually as contagion-related guidelines in which to avoid and health promotion guidelines to emphasize. In the left column, we identified four contagion-related guidelines for what should be avoided in reporting. These include: 1) Avoiding sensationalizing, 2) Not depicting suicide as a crime, 3) Omitting unnecessary details, and 4) Avoiding stigmatized language. Sensationalism refers to highlighting the suicide in an attention-grabbing manner that stimulates the reader, describing suicide as inexplicable and even not preventable, discussing increases in suicide rates as a disease that cannot be stopped, and making suicide seem more commonplace than reality. Next, depicting suicide as a crime includes articles written in the same manner as police report and interviewing police about a suicide. These both treat suicide as criminality and convey judgment [16,92]. Also, it is important to exclude details about the individual who attempted or died by suicide so that a reader does not identify with that individual, making him or her at higher risk through contagion [82]. Not including a suicide note, discussing means, or using photos of where an individual died, their grieving family, or memorials all exclude unnecessary and potentially harmful details. Specifically, including the means in which suicide was carried out is harmful because it can give vulnerable individuals ideas about how to take their own lives [59,92]. Research shows an increase in a particular method of suicide after reporting on a suicide with that method [49,93].

A unique aspect of our study is that we evaluated stigmatizing language and incorporated research in this area because stigmatizing language as this has been associated with contagion-related effects [16,61]. Several aspects of stigmatizing language were included in the World Health Organization's guidelines, these guidelines are often omitted [10,11]. However, stigmatizing language is crucial to suicide prevention since there is evidence that stigmatizing language impacts help-seeking behavior [61]. For instance, highly stigmatizing reports can impact help-seeking behavior and imitative suicides depending on whether the author presents suicide as the only solution to a personal crisis or emphasizes professional help [61]. Stigmatizing language

Table 1. Synthesized Suicide Prevention Safe Messaging Guidelines

Contagion-Related Guidelines	Health Promotion Guidelines
Avoid Sensationalizing:	**Prevention:**
1. **Headlines: Minimize suicide's prominence and do not attempt to grab readers' attention through referring to suicide in the headline; instead use information headlines [86,88]**	1. **Emphasize Prevention**: Present suicide as preventable [76,86,88]
2. **Depicting Suicide Rates: Do not use stronger terms, like epidemic or skyrocketing to describe suicide rates, instead use non-sensationalizing words, such as "rise" or "higher" [86,88]**	2. **Prevention Experts:** Seek advice from suicide prevention experts [86,88]
3. **Depicting Suicide: Do not describe suicide as inexplicable, shocking, or happening without warning [76,86,88]**	3. **Warning Signs:** Discuss the warning signs of suicide [76,86,88]
4. **Normalizing: Do not normalize suicide by presenting it as a common event [76,86]**	
Do Not Depict Suicide as a Crime:	**Treat as a Public Health Issue:**
5. **Police Report Format: Do not write the article like a police report [88]**	4. **Public Heath**: Report on suicide as a public health issue, facts about suicide (not as a crime) [86,88]
6. **Interviewing Police: Do not quote or interview police/first responders about the cause of suicide or attempted suicide [88]**	
Omit Unnecessary Details:	
7. **Suicide Note: Do not include a detailed suicide note [76,86,88]**	
8. **Suicide Means: Do not discuss means of suicide [76,86,88]**	
9. **Unsafe Photos: Do not use photos of location, method of death, grieving family or friends, memorials or funerals [76,86,88]**	
Avoid Stigmatizing language:	**Provide Resources for Getting Help:**
10. **Derogatory/Stigmatizing: Do not use stigmatizing or derogatory language about mental health/illness [88]**	5. **Resources**: Emphasize help-seeking [76,86]
11. **Suicide as Success: Do not referring to suicide as successful/failed attempt [86,88]**	6. **Treatment**: Highlight effective treatments for underlying mental health problems [76]
12. **Commit: Do not use the word "commit" in reference to suicide [86]**	

refers to discussing suicide in a manner that conveys judgment, such as using derogatory or stigmatizing language about mental illness, referring to suicide attempts that lead to death as "successful," using the word "commit." Specifically, the term "commit" is stigmatizing because the other contexts in which this word is used imply judgement, since it is most commonly used in the context of committing a crime [92]. Due to the implicit stigma of the word "commit," "died

by suicide" is the preferred terminology. Comprehensively evaluating stigmatizing language in the existing suicide prevention safe messaging guidelines was a needed step and a unique contribution of this research.

3.2 Health Promotion Guidelines

In the right column of **Table 1**, we identified health promotion guidelines to which news media reports should adhere. These include: 1) Emphasizing prevention, 2) Treating suicide as a public health issue, and 3) Providing resources for getting help. Attitudes toward mental health have been shown to be predictive of help-seeking, stressing the significance of encouraging positive attitudes toward both mental health and utilizing available resources [74]. Emphasizing prevention can occur in several ways, such as educating the community about warning signs for suicide in order to recognize those at risk [92]. Interviewing suicide prevention experts also educates the public on the preventable nature of suicide and combats the suicide myth that once an individual has made up his or her mind, there is nothing anyone can do to prevent their suicide [92]. Next, it is important to treat suicide as public health issue and not as a crime [92,94]. Treating suicide as a public health issue is destigmatizing through presenting mental illness in a similar manner to physical illness. Lastly, it is essential to not only emphasize help-seeking, but also to provide specific resources for where to find help [82,92]. Since research shows that 76% of those who die by suicide have not visited a mental health professional in the last 30 days before their death, there is clearly room to advocate for help-seeking in individuals with suicidal thoughts [1]. While emphasizing help-seeking, it is also recommended to note that there are effective treatments for those with suicidal thoughts, anxiety, or depression [82]. In summary, it is necessary to not only avoid contagion-related reporting, but also promote healthy behaviors. In our methods section below, we describe how we used the framework shown in **Table 1** to qualitative analyze a data set of news media reports on the BWC. **Table 1** also shows our synthesis of safe messaging guidelines.

4 METHODS

As discussed above, we first synthesized the evidence-based safe messaging guidelines related to suicide prevention to create a conceptually driven framework divided into contagion-related and health promotion safe messaging guidelines. Next, we qualitatively analyzed 150 news articles related to the BWC based on this framework. We also conducted a grounded thematic analysis of the news articles to identify emergent themes not included in suicide prevention safe messaging guidelines that surfaced from the data.

4.1 News Article Data Collection

To compile a subset of news articles related to the Blue Whale Challenge, we searched LexisNexis [109], which is a comprehensive online database for digital and printed news articles, using the term "Blue Whale Challenge." All articles included in our dataset were disseminated through online news sources, as well as some through printed media. Since the Blue Whale Challenge first emerged in 2017, we limited our search to articles published from January of 2017 to December of 2018. The articles from earlier articles were mainly discussing endangerment of blue whales, making them irrelevant to the current study. After applying this inclusion criteria and deleting all duplicates from the downloaded list of articles into an Excel spreadsheet, there were 2,165 total articles. Next, we randomly selected 150 articles using the random number generator function in Excel. We excluded any articles that did not pertain to the BWC. Many articles that were selected included a link to an article about the BWC, but were not at the challenge themselves, thus they were removed. Then, we performed another random selection of articles (N=150 articles minus irrelevant articles) to replace those that were deemed were irrelevant. We continued this process

iteratively until we had a data set totaling 150 relevant articles. We reviewed approximately 250 articles to arrive at this number. Assuming that this ratio of irrelevant to relevant articles (approximately 60%) was consistent throughout the data set, we estimate that our data set represented a random sub-sample of about 12% of the relevant articles from our original Lexis Nexis search.

4.2 Data Analysis Approach

We used a hybrid deductive and inductive coding approach [11] to qualitatively code the 150 articles. First, we used the guidelines presented in **Table 1** as a framework for evaluating whether the news articles adhered to or violated the contagion-related and health promotion safe messaging guidelines for suicide prevention. All guidelines were coded as either "Present" or "Absent." Main codes such as "Emphasize Prevention," and "Emphasize Help-Seeking" had several sub-codes, such as emphasizing prevention through education or increased parental involvement. Additionally, we collected demographic information on participants mentioned in the articles, including country of origin, race/ethnicity, sex, and age of the participant. The two researchers who independently coded the articles were undergraduate psychology students trained by the senior researchers who co-authored this work. One researcher is a clinical psychologist who specializes in suicide prevention who trained the students how to interpret and apply the safe messaging guidelines for suicide prevention. The other is a Human-Computer Interaction researcher who specializes conducting in qualitative research guided the data analysis.

After the students coded the data based on the safe messaging guidelines, interrater reliability was assessed using Cohen's kappa [28] and mostly ranged from 83.3% to 99.3% for main codes and 70% to 80.7% for sub-categories within codes. Two low Cohen's kappas (K < 0.6) reflected codes that were not commonly present in the articles. For example, "Photo of location, method, grief, or funeral" was only present in one article. Thus, there was a high probability that agreement would occur by chance due to the low base rate occurrence of this code. We used a consensus building process to resolve any conflicting codes.

Second, we conducted a thematic analysis [18] to identify emergent themes from the news articles. While a number of interesting insights emerged, the two consistent themes we identified were that: 1) Many of the news reports included seemingly factual information (e.g., reported unequivocally as fact) without confirmatory evidence of the facts. 2) News articles often reported on a victim's motivation for participating in the BWC without corroborating this motivation with reputable sources (e.g., the victim or their families), and we discuss these two themes in more detail in our results.

The 150 articles were coded in increments of 50 to assess whether qualitative saturation had been reached in two different regards: First, we assessed whether the proportion of articles that adhered to or violated the safe messaging guidelines changed significantly from one batch of articles to the next. Second, we assessed whether theoretical saturation of our emergent themes was reached. After comparing the three increments (i.e., 1-50, 51-100, 101-150), we confirmed that proportions across the subsets remained stable and that no new themes emerged from the addition of 50 new articles. Therefore, the 150 articles were sufficient to conclude our analysis and report our findings.

5 RESULTS

The goal of the current study was to evaluate the adherence to safe messaging guidelines through excluding contagion-related messages (e.g. use of sensationalizing headlines) and including health promotion messages (e.g. emphasize help-seeking). Overall, we found that there were frequent contagion-related violations and shallow adherence to health promotion messaging guidelines. Prior to reporting our results, we need to acknowledge that there are ethical considerations for conducting suicide related research. As researchers, it is difficult to both robustly examine how others report on suicide-related topics without doing the same ourselves. Thus, we did our best

attempt to remove unnecessary details, while providing enough context to show how the articles in our dataset violated safe messaging guidelines.

5.1 Descriptive Characteristics of the News Articles

When present in the article, we coded for the country of origin, race/ethnicity, sex, and age of the participant. Most participants were from India (79), but other articles came from Argentina (12), Egypt (4), and United States (2). The most common race was Asian (84%), followed by the ethnicity Hispanic (14%). Additionally, more males were listed as participants of the challenge, with 64% males and 36% females. The mean age of participants was 16.5 with a standard deviation of 4.2. While our article search was world-wide, the large majority (n = 120) of the articles were from Indian news sources. Most articles were written as a public service announcement that also mentioned one to three individuals who had participated in the challenge. Participants nor their families were rarely interviewed for the news article. Interestingly, very few articles interviewed the participants about their participation in the challenge and personal characteristics. In order to determine the credibility of these sources, we also examined the most read news sources in the respective country of the article. Half of the articles from India were written by the top twenty most popular news sources [9], while several of these sources were rated by readers as "somewhat trustworthy" [9].

5.2 Contagion-Related Guideline Violations (RQ1)

Overall, 81% of articles violated at least one contagion-related guideline. Most articles violated one to four guidelines (68%), with violating more than four only occurring a few times (<1%). The most common violations were normalizing suicide, sensationalizing headlines, discussion of means of suicide, and use of the word "commit."

5.2.1 Avoid Sensationalizing

Sensationalizing headlines refer to headlines written in a way that either highlighted the suicide, death, or self-harm in an alarming and attention-grabbing manner. For this study, 31% of the articles examined had sensationalizing headlines. For example, one title reads, *"12-year-old Egyptian schoolboy commits suicide to fulfill 'Blue Whale' challenge"* [34]. This title both sensationalized death by suicide and used stigmatizing language when invoking the word "commit." Another headline read, *"ALERT: Parents of Dead Teen Find Clues to Her Suicide on Bedroom Wall- It's a Deadly Game, and its Spreading"* [97]. Again, this headline sensationalized suicide through attempting to grab readers' attention rather than using a more balanced approach that emphasized prevention.

Discussing suicide rates with words such as "skyrocketing" or "epidemic" does not present suicide in an accurate manner, such as presenting statistics on suicide rates. Fortuitously, this type of language was only present in two articles. One article used the term *"surge,"* which implies a very sudden, quickly rising, and inexplicable change while depicting a picture of a the individual's means of suicide [20].

Presenting suicide as inexplicable, shocking, or happening without warning is problematic because it presents suicide as non-preventable and random, when in reality, there are warning signs to suicide. Five articles presented suicide in this manner. One article writes, *"We were shocked, we thought our child was normal, we could never imagine he was going through so much"* [98]. Unfortunately, this article went on to include a suicide note, include a drawing of the child's plan to end his life, write the article like a police report, and use a sensationalizing headline [98].

Normalizing suicide by presenting it as a common event both oversimplifies suicide and makes it seem like a viable option to stress and hardship [82]. This occurred in 33% of articles, in which

many stated large numbers of BWC related suicides, such as claiming the BWC had taken over a hundred suicides in Russia alone over the last two years. Statements like these depicted suicide as a common event happening frequently among youth. One article simplifies suicide in such a way that it attributes suicide due to lower parental involvement. While there may be an aspect of truth in that lower parental involvement can increase risk for suicide, the following quote presents suicide as the natural and reasonable response to inadequate parenting. The death of this girl *"is a wake up call for all the parents. Today parents don't have time to interact with their children and often they are busy with their professional lives"* [99]. This framing presented suicide as a normal and viable solution to a problem, which has the potential to encourage suicidal behaviors in others [82].

5.2.2 Do Not Depict Suicide as a Crime

Reporting on suicide as a crime refers to writing an article with only detail about the individual and low emphasis on help-seeking or prevention. While only 15% of articles quoted or interviewed police or first responders, 31% reported on suicide like a crime. One article reads, *"The victim [girl who attempted suicide] has been identified as ---------- of ------- village"* [100]. This article simply identifies a girl who attempted suicide, discusses her means of suicide, includes a vivid picture depicting self-harm, reports on other incidents in the area, and mentions a special crime branch of the government. Note that we do not include names or locations in order to protect individual's privacy.

5.2.3 Omit Unnecessary Details

A detailed suicide note includes a letter with a vivid description of an individual with suicidal thoughts. Only 7% of the articles included a suicide note. For example, one reporter quotes a suicide note saying, *"I don't deserve to live', 'No one cares about me', No one loves me', 'They just use me', 'I should just die'"* [98]. The level of description in the individual's thoughts has the potential to lead someone else with similar thoughts to consider suicide [82,94]. This article not only presents dangerous information, but also fails to mention help-seeking behaviors for individuals distressed by this note and suicidal thoughts.

Next, discussing individual means is a form of providing too much detail about a suicide and may even give vulnerable readers ideas if they are considering suicide. Articles frequently described the type of weapon or item used to die by or attempt suicide (32%), occasionally even providing images of the type of item used. Several articles even mentioned the family of medicine that an individual consumed to attempt suicide. Providing both information about and pictures of means can give viewers ideas about how to end their own lives [49,93]. While this was not specific to the safe messaging guidelines, we also found that in addition to including details about a suicide, journalists typically included detailed information on how individuals became involved with the challenge and the steps of the challenge. Since means of suicide ought to be omitted, the steps of the challenge listed could also serve as a guide for how vulnerable readers could participate in the challenge. This is a unique aspect of reporting regarding the BWC because there could be many unnecessary details given regarding an individual, their suicide or thoughts, and the BWC itself.

The last guideline says do not include photos of location, method of death, grieving family or friends, memorials or funerals. This was one of two guidelines that not a single article violated.

5.2.4 Avoid Stigmatizing Language

The next guideline regarding stigmatizing language is do not use derogatory or stigma-invoking language about mental health or illness, which occurred in 13% of the articles. For example, an individual refers to self-harm as *"satanic stuff"* [101] and another claims that individuals self-harm and attempt suicide because they *"crave attention and expect to be treated as*

individuals instead of spoiled brats" [102]. Again, this language implies judgment and trivializes mental illness.

Another guideline regarding stigmatizing language is do not refer to suicide as a successful or failed attempt because suicide should never be equated with success. This was the other guideline that no article violated. While most articles included at least one aspect of stigmatizing language, most used the word "commit" when referring to suicide as opposed to "successful suicide" or "failed suicide." The word "commit" was used in 58% of the articles. The use of this word implied judgment and wrongdoing, which are not beneficial to those struggling with mental illness [16,92]. Again, stigmatizing language has the potential to decrease help-seeking behavior, which is the opposite of what news reporters should be urging readers to do.

5.3 Health Promotion Guideline Adherence (RQ2)

Overall, 91% of the articles adhered to at least one health-promotion guideline. Most articles adhered to one to four guidelines (85%), with zero (9%) or five or more occurring infrequently (5%). The most commonly adhered to were emphasizing prevention and reporting on suicide as a public health issue.

5.3.1 Emphasize Prevention

Suicide prevention is a term that encompasses anything a person does for another or him or herself that inhibits acting on suicidal thoughts. Fortunately, 79% of articles emphasized prevention of one or more types. Most articles targeted prevention through parental involvement and education. For example, one author reports, *"Even countries in which no incidents related to the challenge have been reported are making efforts to educate parents"* [29]. Additionally, articles attempted to educate teachers and school counselors to look out for signs that their students were playing the game. *"Even in schools, teachers and counsellors should handle this matter with sensitivity"* [22]. While parental and teacher prevention are important, very few articles mentioned teaching children online safety skills or addressed peer suicide prevention programming.

Only 14% articles interviewed a mental health professional or any sort of doctor about the issue. This was a similar rate to interviewing the police about a suicide (15%). One article both interviewed an expert, highlighted the complexity of suicide, and emphasized treatment by saying, "*Suicide, experts say, is rarely connected to a single event or cause, with some 90 percent of those who take their own lives already suffering from mental illness, often untreated depression*" [29]. Interviewing experts is essential as these individuals are more likely to understand safe messaging and how to promote help-seeking and prevention through the media.

Next, 20% mentioned at least one warning sign of suicide. The majority noted that people should watch for general changes in behavior but did not list a specific warning signs or what types of behavioral change could be a risk for suicide. For example, "*She said the teenager started behaving indifferently and was always engaged in weird activities*" [103]. While a sizeable portion of authors included a general statement to look for changes in behavior, there is room for improvement in how frequently this information is included and the detail of warning signs included.

5.3.2 Treat as Public Health Issue

Instead of reporting on suicide as a crime, it is beneficial to report on suicide as a public health issue. This means treating suicide as a disease and problem that can be addressed on a societal level, as opposed to personal weakness that an individual ought to deal with on his or her own. Reporting on suicide as a public health issue was the second most commonly adhered to health

promotion guideline (79%), yet presented itself differently due to the nature of the BWC. Most authors took a perspective that the government needed to ban or take down the game because of the public health issue of suicide supposedly resulting from the game. For example, one article reads, "*The federal government has directed technology giants, including Google, Facebook, Microsoft, and Yahoo to remove all links which direct users to the dangerous online game 'The Blue Whale Challenge'*" [103]. While this is not a traditional type of public health reporting, addressing the problem of suicide from a governmental perspective highlights the societal level intervention due to the nature of the BWC.

5.3.3 Emphasize Help-Seeking

Unfortunately, emphasizing help-seeking was present in only 37% of the articles. Noteworthy, the source of help was different due to the nature of the BWC. Overall, 18% of the articles suggested contacting the police as a form of help-seeking through a suggestion such as, "If need be, parents should inform the police to ensure the link between the child and the game's administrators is severed" [22]. Only 7% of the articles that emphasized help-seeking suggested seeking mental health treatment. One of the few articles that emphasized professional psychological help wrote, "The intensity of the problem could be successfully treated by the professional psychologists and counsellors, before this could become a serious mental health issue" [33]. Additionally, only three articles had direct resources available, such as a suicide prevention hotline. It is recommended that when speaking about a topic as sensitive and potentially distressing as suicide, one should provide information on where to get help immediately. While emphasizing any form of help-seeking is beneficial, emphasizing psychological help and providing specific resources is essential and has room for improvement.

Lastly, highlighting effective treatments for underlying mental health problems is an important aspect of encouraging help-seeking in order to show the public there are treatments that have been shown to be effective. Only eight articles spoke about mental health treatment. One article both reports on suicide as a public health issue and highlights treatment with the following statement, "Mental health issues are just like cardiac problems, diabetes, hypertension or other illnesses. Recognize it for what it is, treat it, change your lifestyle as required, and provide the support needed to improve the condition. When in doubt, seek professional help" [14]. In sum, while many articles did in fact emphasize prevention and report on suicide like a public health issue, there is a need for a variety of types of prevention and help-seeking.

5.4 Emergent Themes Regarding Misinformation in BWC News Reporting (RQ3)

The two themes that emerged from out dataset included problematic reporting regarding the overuse of casual language and reasons associated with participation without corroborating evidence.

5.4.1 Causal Language Used within News Media Reports

One of the most prevalent themes and areas of concern was use of causational and confirmatory language used by the authors. The authors of the news articles often included one equivocating word like "allegedly" or "supposedly" connecting the BWC to a suicide, but then proceeded to report on the death as if it were undoubtedly caused by the challenge with little or no confirmatory evidence. This was present in 81% of the articles. For example, one article begins by saying that "*a 16-year-old boy from ----- is the fifth victim of notorious Blue Whale challenge in India*" without providing any evidence of the attempted suicide being linked to the challenge or fact-based information about the previous suicides. The article went on to contextualize the challenge with additional unconfirmed details, "*Starting from Russia, the Blue Whale game or Blue Whale Challenge has finally arrived in India via Brazil, China, Italy, Argentina, Spain, Venezuela, Georgia and other European countries*" [104]. While there is some evidence that the challenge may

have started in Russia, some studies also found evidence that the challenge may be present in Italy [53,110]. Further, stating that challenge is widespread without proper evidence could be seen as sensationalizing, spreading panic, and potentially disseminating misinformation since the reach of the challenge cannot be determined.

Another article with the headline of "*Class 10 Panchkula Student Found Hanging, Family Blames Blue Whale*" [98] is another salient example. This article used a sensationalizing headline, named the BWC as the means of suicide, included unnecessary personal details about the participant, quoted a suicide note, and used stigmatizing language. In addition, the causal link between suicide and the BWC was based only on imagery drawn by the child prior to their death. Yet, the article states, "*they did not find any trace of the blue whale game on his phone or laptop,*" but still speaks of the death only in relation to the challenge. The article also stated that the police had not yet investigated the matter, and the post-mortem had not been released. In this case, not only was there no concrete evidence relating to the challenge, the authors stated that there was a lack of evidence, despite attributing the death to the BWC.

Further, 21% of the articles said that those who tried to quit playing the game received threats to leak personal information or threats against their families, such as "*if they fail to do so [complete the given tasks], they are threatened of untoward incidents or harm to their family members or themselves*" [104]. This article made the claim that all participants are threatened if they do not complete a task. However, this claim was not verified by interviews, citation of evidence wherein this occurred, or any other form of support. Another article referred to an image on Twitter saying "*players of the challenge can't stop once they've started; they are blackmailed and cyber-bullied into completing the 'game'*" [105] with a similar lack of confirmatory evidence. Overall, journalists provided little, if any, factual information to back up their often causal and sensationalized claims, which contributed to the spread of misinformation about the BWC. Using causal language without evidence reveals concerns of misinformation and points to a larger problem related to the spread of potentially harmful online information.

5.4.2 Unconfirmed Reasons for Participation

Our thematic analysis also revealed that reporters commonly (~50%) spoke of why participants participated in the BWC. Upon the emergence of the theme of reported reason(s) for participation, we coded for these reasons given. The most common reason reporters gave for participation in the BWC was game addiction (23%). One journalist explained, for instance, "*[he] was addicted to some computer games for the past few months and he used to always play some games on his mobile and laptop and used to sleep in the morning after playing computer games the whole night*" [106]. The journalist attributed this explanation based on a quotation from a previous interview with a different media source. Another article cites the BWC as common in a particular geographical location because "*addiction to online games is extremely high among children and teenagers because it offers an escape from the dreary reality of a small town*" [107]. Yet, the author reported no evidence that either the BWC or internet addiction was high in that area. Instead, journalists speculated on why adolescents would participate in this challenge tying it to the popular narrative of online gaming addiction among youth. Indeed, online game addiction has become an issue of interest for youth [50]. Yet, the reporters rarely interviewed the individual who had participated in the challenge to support this claim, though they occasionally interviewed family members or police officers when making these claims.

Only 8% of the stated motivations to participate included depression or previous suicidal ideation. An interview with doctors who treated a girl who attempted suicide while supposedly playing the game states "*the girl is suffering from severe depression,*" as this was her second suicide

attempt in a short period of time [108]. However, the participant was not interviewed about the challenge or her history. Additionally, instead of focusing on mental health and suicide prevention, this article focused almost exclusively on the BWC. The author was more concerned with publishing a detailed report on the dangers of the BWC, using the girl as an example, rather than emphasizing prevention, help-seeking, or providing resources. We found one article that even pointed out this problem—that depression and mental health issues were being "*sidelined*" and attention was given to the challenge instead of including reliable information and resources on suicide prevention [96]. The author referred to media reporting on the subject as "reckless journalism" and panic spreading due to conjecture and speculation in the face of opposing evidence [96]. Again, listing reasons for participation with little or no evidence reveals an issue in both journalism and online sharing of information, the sharing of potentially harmful misinformation.

6 DISCUSSION

We evaluated 150 newspaper articles to assess whether they adhered to or violated suicide prevention safe messaging guidelines. Overall, 81% of the articles violated at least one contagion-related guideline; fortunately, 91% adhered to at least one health promotion guideline, albeit on a shallow level. Thematic analysis also revealed concerns regarding the dissemination of potentially harmful misinformation within BWC news reporting. We discuss the implications of our results and make recommendations for research, practice, and policies that can potentially mitigate harm caused by irresponsible reporting on viral self-harm social media challenges.

6.1 BWC News Reporting May Perpetuate Suicide Contagion

A unique characteristic of the Blue Whale Challenge is its dual potential for propagating suicide contagion effects, through encountering the challenge itself via social media, exposure to violations of suicide prevention reporting guidelines, and consuming (mis)information about the challenge via news media reports. As such, avoiding contagion-promoting content, dispelling myths, and providing factual information through trusted news sources becomes paramount, as the widespread concern over users interacting with false content via social media surges [3]. Yet, our analysis found that news media reports on the BWC violated several contagion-related safe messaging guidelines, most commonly by normalizing suicide, using sensationalized headlines, discussing of means of suicide, and using of the word "commit" in relation to the suicide challenge. Therefore, we suspect that news media reports about the BWC, unfortunately, may have contributed to rather than detracted from the potential contagion effects associated with the challenge. Additionally, our findings regarding the discussion of means of suicide is noteworthy. The BWC is by definition of the challenge a means to suicide; therefore, it is debatable as to whether reporting on the BWC itself is an inherent violation of suicide prevention safe messaging guidelines.

Through synthesizing multiple safe messaging guidelines, we found that there is room for improvement in these guidelines for media reporting regarding stigma, considering most articles contained at least one aspect of stigmatizing language. Thus, a contribution of our work was that we incorporated stigmatizing language from the World Health Organization's safe messaging guidelines [57,87] into our synthesized framework. Presenting other options to managing distress besides suicide has the ability to increase help-seeking behavior, whereas stigmatizing language has the ability to decrease help-seeking behavior [61]. Therefore, avoiding stigmatizing language should be included in all future safe messaging guidelines. Our findings regarding stigma can thus inform practice, such as educating reporters about why stigmatizing language can be harmful.

We found notable similarities and differences when comparing our findings to existing work on similar topics. For instance, studies have examined the adherence to or violation of safe messaging guidelines as they related to news media reports on other trending topics, such as the

death of Robin Williams [22]; these researchers found that most of the contagion-related guidelines were followed, while health promotion guidelines (e.g., help-seeking) were less often implemented [22]. We made similar conclusions about the lack of health promotion. Yet, some key differences between this prior work and our own include that we included stigmatizing language when conducting our analyses and that we found significantly more violations of the safe messaging guidelines when applied to news media reports of the BWC. Another key difference is that news reports on Robin William's death were unequivocally about an individuals' suicide, rather than a technology-mediated phenomenon like the BWC. While we cannot confirm causality, we do ponder whether the BWC, being at the unique intersection of (sometimes elusive reports of) social media behaviors and suicide, may have contributed to the sensationalized reporting on the phenomenon. It is possible that the virality and "game-based" nature of the BWC decreased reporters' sense of urgency and duty to report on the dire topic of suicide in a responsible manner. On the contrary, it is also possible that general news media reporting on suicide is equally as irresponsible. Therefore, future research might consider comparing news media reports on suicides associated with the BWC to other news media reports on suicide to disentangle potential interaction effects between suicide reporting and social media phenomena.

The combination of self-harming content with technology is of particular concern. Previous human-computer interaction researchers have suggested that technology mediated self-harm has the potential for more damage than the same behavior without technology due to the size of the audience and availability of the content [68]. Digital self-harm and suicide is thus essential to monitor due to the potential scope of harm. Next, we discuss the implications of our findings regarding BWC news reporting and health promotion.

6.2 Meaningfully Promoting Health and Online Safety of Youth

A finding that previous studies and ours have in common is the general lack of resources authors publish alongside their reports that relate to suicide prevention and health promotion [6,14]. While most articles had at least one aspect of health promotion reporting, the overall depth was lacking. For instance, articles would emphasize that parents should pay closer attention to their children's on and offline activity as a means to prevent participation in the BWC, but this is where the majority of the articles stopped. As such, the articles merely scratched the surface when touching on the topics of suicide prevention and adolescent online safety. Ironically, the suicide prevention safe messaging guideline that were violated often occurred simultaneously within articles that were trying to raise awareness, educate parents, or promote prevention of the BWC. This is similar to previous research that shows public awareness announcements regarding the BWC are likely to elicit comments asking how one can participate on social media. Social media users are thus having the opposite of their intended effect, contributing to contagion when they are trying to promote awareness [46]. Prior research highlights the importance of carefully studying the effectiveness of suicide prevention campaigns and public service announcement and found that some interventions may cause more harm than good [48]. We believe that this may also be the case given the imbalance of contagion related violations versus health promotion guidelines adhered to within the news coverage of the BWC articles.

Instead of encouraging parents to monitor their children or restrict their internet usage, as suggested by many of the articles, research has shown the importance of using resilience-based approaches to adolescent online safety that help prevent harm from online risk exposure [90]. In contrast, sensationalized news reports on the inherent risks of the BWC are likely to promote parental fear and control, rather than youth resilience and self-regulation [88]. News reporting on the BWC, suicide, and self-harm could all be opportunities for parents to have an open conversation with their children about mental health and suicide, if these reports were written responsibly and factually, omitting unnecessary details. While the topic of suicide is a sensitive one, especially for children, there are specific recommendations for talking to youth about suicide

[19]. These recommendations include acknowledging the serious nature of suicide, helping children process their feelings about the issue, conveying risk factors, and telling children how to find help for others they know who might be contemplating suicide [19]. These guidelines give parents guidance on how to address an important subject with their children in a safe manner, so that children are not first exposed to suicide either alone or in a non-supportive environment but exposed in the home with the support of their parents. Thus, news articles could link to resources such as these within the articles themselves, rather than making ambiguous suggestions for parents to talk with their children. While help-seeking was often encouraged, such as contacting the police or speaking with an adult, publicly available resources, such as a suicide prevention hotline [73] or crisis text line [65] were not included in many of the articles. Providing a crisis hotline has been shown to reduce symptoms of depression and anxiety and this information should be provided when discussing sensitive subjects such as suicide [73]. Human-computer interaction researchers suggest including resources and help-seeking information when a certain keyword is searched in relation to any type of digital self-harm [64]. These recommendations and interventions are typically for online self-harm communities, as opposed to news reporting or exposure to social media challenges that only include the participant and curator. While this is a distinctive type of digital self-harm promotion, similar resource linking could be provided based on search terms as previously suggested. Yet, the distinctive aspect of the BWC is noteworthy and future research efforts ought to examine digital self-harm and harmful social media challenges in a non-community-based setting.

Through educating parents on the presence of online risk and equipping them with the tools to address both internet safety and suicide with their children, we can deepen both the prevention of suicide and promotion of healthy behaviors. There are several simple opportunities for journalists to include evidence-based resources in their articles to promote suicide prevention and mental health. In the next section, we discuss the implications of our findings as they relate to the dissemination of misinformation online.

6.3 Misinformation in the Era of Viral Social Media Challenges

The emerging themes that arose from our thematic analysis were regarding misinformation due to the lack of evidence for claims linking suicides to the BWC and reasons for participating in the BWC. Most articles contained causational language (81%) without confirmatory evidence that the BWC was indeed related to the incident being reported. While some of the articles included an interview, most news articles speculated or made assumptions. This is concerning due to the potential for suicide contagion effects given the nature of the challenge. Even if a death is not confirmed as being related to the BWC, news reports making causal inferences to this effect could potentially create imitative suicides as research shows a parallel between media coverage of suicides and suicide rates [38,39,41,70]. As such, inaccurate or unconfirmed reporting on the BWC is a uniquely harmful form of online misinformation. Further, many journalists speculated about the reasons that adolescents participated in the challenge, such as having a gaming addiction. Meanwhile, Lanette et al. [51] show that addiction narratives are often compelling but largely unfounded. Therefore, inferring that game and online addictions are the primary motivations for participating in the BWC could unintentionally lead to interventions that do not target the real problem. Thus, more research needs to be done to truly understand the motivations behind why individuals might consider participating in the BWC and other viral social media challenges.

Overall, we found evidence that journalists were not only violating suicide prevention safe messaging guidelines when reporting on the BWC, they were also potentially disseminating misinformation about the challenge. Starbird et al. [80] explains how reporters are likely to contribute to the dissemination of misinformation online. Journalists play a unique and key role since they are more likely to be the first to create original content and have their posts shared more frequently more frequently by others [80]. During crisis events, they often publish first and later verify the information, which may have already been shared broadly by others [12]. As such, there is a trade-off between balancing the time frame in which journalists release information

and the verification of information, which leans toward a speedier time frame over verification in times of crisis [80]. Similarly, most of the news stories we examined provided little or no evidence that suicide was definitively a result of the BWC, nor did they corroborate the reasons why they made the connection through eyewitness account from the victim or from friends and family members. Thus, some of the ambiguity surrounding the BWC can be attributed both to difficulty of verification as previously mentioned and to lack of verification before publishing. Additionally, concerns arise from the accessibility of online information that is unverified and able to be propagated via social media [8,80]. The key implication that arises from our two emergent themes is that harm can be done from poor quality news reporting whether or not the story is valid (and especially if it is invalid). Since online news articles can be shared at a rate higher than ever before, how we approach suicide prevention safe messaging and misinformation needs to be adapted. Otherwise, these articles could put vulnerable individuals in harm's way unnecessarily. It is noteworthy that this information is both online and has high potential for harm.

Sharing cascades of digital content is the key factor in the spread of misinformation [75]. Recent studies [23,40] have identified three major concerns when consumers use information obtained from online sources: (1) Misinformation has the potential to become viral just like reliable news; (2) Social media contains information contradicting reference standards/guidelines; and (3) Social media has the potential to change the beliefs of information consumers about controversial topics such as vaccinations [23,52]. Though such information is predominantly created by fake news websites, individual hoaxers, and other propaganda sources, our study provides evidence of potential misinformation in reporting, where the articles unequivocally attributed many suicides to the BWC with little or no evidence. Given the importance of ensuring that the reliability and accuracy of information used while understanding crisis situations, it is crucial that we develop evidence-based practices, policies, and tools to minimize the spread of misinformation and suicide contagion effects associated with portrayal of self-harm as it relates to suicide, the BWC, presentations of self-harm on social media, as discussed below.

6.4 Implications For Practices, Policies, and Tools

A multi-pronged approach is necessary to improve news reporting on suicide, self-harm, and the BWC in a way that decreases both potential suicide contagion effects and the dissemination of information. In terms of practice, we first recommend widespread training to educate news reporters on how to responsibly report on self-harm and suicide. News outlets could require reporters to take an online course on the suicide prevention safe messaging guidelines, as suggested and available through the Reporting on Suicide website [94], which could potentially provide better guidance to reporters on developing content adhering to safe messaging guidelines. This training would be particularly salient for reporters who typically cover social media phenomenon, rather than mental health and suicide given that they are likely less attuned to the importance and gravity of the intersecting topics of social media, suicide, and self-harm. From a policy perspective, one issue in safe messaging is the variety of overlapping guidelines that we identified, thus making it difficult for reporters to choose which to follow. To this point, we synthesized the guidelines to form a more comprehensive framework of suicide prevention safe messaging guidelines, which helped facilitate our own analyses. As such, we urge the various authorities on suicide prevention safe messaging to come together to form a consensus and a world-wide standard around safe messaging. These guidelines could be modified for different cultural contexts; nonetheless, one standard would help reduce ambiguity and confusion regarding responsible reporting standards. Further, the standards need to be updated due to the advent of social media and the potential viral nature of self-harm challenges, like the BWC.

Since journalists play a key role in dissemination of information, suicide prevention safe messaging guidelines also ought to include recommendations on not only how to report on

suicide, but how to do so credibly. A key issue is that the sources used in news media production constantly change, compounded by the constraint that only minimal barriers are applied to a developing news story [2,3,45]. For example, reporters should leverage resources that include tips on what makes a story non-credible, the categories of non-credible news, and tips for how to avoid writing in a way that makes an article untrustworthy [111]. In short, journalists should be strongly discouraged to publish unverified claims as to prevent the spread of misinformation and suicide contagion effects. News reporters may hear a piece of information and not question that information, yet more validation and confirmation are needed before reporting on something as serious and potentially harmful as suicide.

Prior work has suggested some approaches for reducing misinformation in the news, which include: (1) providing feedback from credible sources to users that a particular information may be non-credible, (2) developing and implementing algorithms for detecting information promoted by bots and (3) identifying the sources of origin of most non-credible news and subsequently reducing the promotion of this information [36,76]. Integrating available verified information may increase the trustworthiness and veracity of the information available, while incorporating features such as crowdsourcing, whereby users are encouraged to report inaccurate and misleading information, is another possible strategy for preventing the spread of misinformation.

Others suggest using hazard mitigation frameworks as a form of Prevention through Design (PtD) [81], which acts as a prioritization schema for dealing with crisis events. This framework consists of three approaches, including: (1) designing the hazard out; (2) guarding against hazard; and (3) warning against the hazard. In the specific case of news media reports on the BWC, we might "design out" the hazard or risk of irresponsible news reporting on BWC by developing natural language processing algorithms [48–52] that identify news media reports on the BWC. Just as human-computer interaction researchers are working on software to aid moderators in identifying harmful content on social media platforms, natural language processing algorithms can be applied to aiding journalists in identifying potentially harmful information in their own work [25]. These automated tools might find ways to suppress or correct these articles based on an evaluation of (1) whether they violate a substantial number of suicide prevention safe messaging guidelines, (2) make unfounded causal claims, or (3) have other features that are highly indicative of spreading misinformation about the BWC. Software has been able to predict suicidal ideation and psychological state among online mental health communities [32]. Since social computing researchers have addressed automated analysis of social media forums and communities for harmful content, we urge similar applications in the field of journalism [53,26]. Similar approaches have been applied in previous research that evaluated bullying and cyberbullying in news media [56]. These approaches can be used to analyze a variety of features, such as emotional tone, causal language, feelings discussed, and specific words like related to death or depressive feelings. Thus, we urge news media platforms to incorporate such algorithms to minimize the spread of potential suicide contagion effects and misinformation. Additionally, in order to warn against the hazard, websites could warn users about sensitive content. Journalists could be asked to inform readers at the beginning of the article that there is sensitive material included. Disclaimers could automatically be included in such articles to warn readers of the nature of the content in the portrayed in an article. Through a multi-faceted, hazard control approach, there are several practical ways to develop tools that may improve safe messaging, decrease the dissemination of misinformation, and curb contagion effects.

6.5 Limitations and Future Research

We recognize that it is extremely difficult to find a balance between research on important problems such as suicide without exacerbating contagion effects ourselves. For this reason, we endeavored to provide the least graphic information as possible while still conveying the reality of poor reporting, answering our research questions, and conducting rigorous research. We also included a warning at the beginning of the paper so that vulnerable readers understand the type of material present in this paper. While we took several precautions to avoid violating safe

messaging guidelines, this is a complex issue that is far from solved. As a research community, we need to be aware of how our research on suicide can affect those within the community and outside the community negatively.

There are several limitations of this study that can inform future research. First, due to the prevalence of BWC reports in India, most of the news articles we analyzed were from this country. Prior research has examined newspaper reporting on mental illness in India and found the manner of reporting to be slightly different from Westernized cultures, as news articles in India are more stigmatizing regarding suicide [81]. Additionally, collectivist cultures, such as Asian countries may be especially susceptible to stigmatizing beliefs due to the value placed on conforming to one's peers [5]. Although India has higher suicide rates than the world average, it is also important to study public health topics like the BWC in other countries [95]. This emphasizes the importance of examining the presentation of mental illness cross-culturally and combatting stigma at a societal and global level. As such, future research should make an attempt to broaden the geographical scope when investigating online self-harm challenges to inform practice and policy tailored to each area.

Another limitation we faced was our inability to consistently record the number of "shares" that each article received. One article was shared 1,100 times, suggesting that it had a fairly broad reach. Yet, we were unable to ascertain how many times each article was read or shared to determine overall viewership. Additionally, we analyzed a subset of 150 randomly selected news articles from Lexis Nexus. We also acknowledge the lack of quantifiable information on the trustworthiness of news sources as a limitation of this research. Many internet users currently receive news reports through a variety of sources that vary in their reliability. Therefore, future research should better assess the quality of news sources and potential for misinformation. Finally, future research could examine suicide clusters after an outbreak of and poor reporting on self-harm and suicides social media challenges, as is common in the contagion literature within offline contexts. By doing this, we may be able to formulate more conclusive patterns on how and whether viral social media challenges pose a threat to our youth.

Widespread access to the Internet and the subsequent explosion of social networks now means that users within these information spaces have massive amounts of data at their fingertips. Consequently, a significant change in the flow and lifecycle of unverified and misleading information has been the result. Specifically, the concept of belief perseverance [7,24], in which individuals hold concepts as true even in the face of overwhelming evidence to the contrary, has greatly accelerated with the advent of social media. To manage this abundance of online information that can be based upon misinformation and speculation, users must rely heavily on their past knowledge, heuristics, and social cues to determine the credibility of information and to shape their beliefs. However, often these methods are inadequate. This filtering and integrating useful, trustworthy, and valid sources of information is a complex cognitive activity that is becoming increasingly difficult with the growing amount of non-credible news available. This work calls for developing and applying systems, tools, and decision aids to minimize contagion effects and enable individuals develop to the situational awareness needed for evaluating the credibility of the information and for incorporation in news aggregation portals and social media in response to the propagation of misinformation.

7 Conclusion

Our study synthesizes suicide prevention safe messaging guidelines and evaluates news practices regarding the Blue Whale Challenge. In sum, it is evident that there is still room for improvement in the realm of reporting on suicide, self-harm, and harmful online challenges in the media to combat the problem of contagion. A major contribution of the current study is the synthesis of what were previously disjointed sets of suicide prevention safe messaging guidelines. Overall, our findings reinforce the notion that adherence to safe messaging guidelines is not yet adequate. We

also found evidence of misinformation, as authors often used causational language and speculated about why individuals participated in the BWC without evidence. Our findings lead to recommending interventions such as further education of news reporters and application of recent advances in natural language processing software in order to decrease contagion. There is an opportunity to dampen the harmful consequences that result from poor quality reports of suicidal behavior and decrease the potential for suicide contagion in reporting on online self-harm and suicide.

RESOURCES FOR SEEKING HELP
If you or a loved one has thoughts of suicide, please contact the following resource:
Suicide Prevention Hotline: 1-800-273-8255

ACKNOWLEDGMENTS

REFERENCES
[1] Brian Ahmedani, Gregory Dimon, Christine Stewart, Arne Beck, Beth Waitzfelder, Rebecca Rossom, Frances Lynch, Ashli Owen-Smith, Enid Hunkeler, Ursula Whiteside, Belinda Operskalski, Justin Coffey, and Leif Solberg. 2014. Health Care Contacts in the Year Before Suicide Death. *Journal of General Internal Medicine* 29, 6: 870–877.

[2] Hunt Allcott and Matthew Gentzkow. 2017. Social Media and Fake News in the 2016 Election. *Journal of Economic Perspectives* 31, 2: 211–236.

[3] Hunt Allcott, Matthew Gentzkow, and Chuan Yu. 2019. Trends in the Diffusion of Misinformation on Social Media. *Research and Politics* 6, 2: 1–8.

[4] Astier Almedom and Douglas Glandon. 2007. Resilience is not the Absence of PTSD any more than Health is the Absence os Disease. *Journal of Loss and Trauma* 12, 2: 127–143.

[5] Soontae An and Hannah Lee. 2018. Suicide Stigma in Online Social Interactions: Impacts of Social Capital and Suicide Literacy. *Health Communication* 34, 11: 1340–1349.

[6] Nazanin Andalibi, Pinar Ozturk, and Andrea Forte. 2017. Sensitive Self-disclosures, Responses, and Social Support on Instagram: The Case of #Depression. *Computer Supported Cooperative Work*.

[7] Craig Anderson. 1981. Abstract and Concrete Data in the Perseverance of Social Theories: When Weak Data Lead to Unshakeable Beliefs. *Journal of Experimental Social Psychology* 19: 93–108.

[8] Monica Anderson and Jingjing Jiang. 2018. *Teens, Social Media, & Technology*. Pew Research Center. Retrieved from https://www.pewinternet.org/2018/05/31/teens-social-media-technology-2018/

[9] Zeenab Aneez, Taberez Neyazi, Antonis Kalogeropoulos, and Rasmus Nielsen. 2019. *India Digital News Report*.

[10] Yasir Arafat, Murad Khan, Thomas Niederkrotenthaler, Michiko Ueda, and Gregory Armstrong. 2019. Assessing the Quality of Media Reporting of Suicide Deaths in Bangladesh Against World Health Organization Guidelines. *Crisis: The Journal of Crisis Intervention and Suicide Prevention*.

[11] Yasir Arafat, Bithika Mali, and Hasina Akter. 2019. Quality of Online News Reporting of Suicidal Behavior in Bangladesh against The World Health Organization Guidelines. *Asian Journal of Psychiatry* 40: 126–129.

[12] Ahmer Arif, Kelley Shanahan, Fang-Ju Chou, Yoanna Dosouto, Kate Starbird, and Emma Spiro. 2016. How Information Snowballs: Exploring the Role of Exposure in Online Rumor Propagation. *Computer Supported Cooperative Work*.

[13] Gregory Armstrong, Lakshmi Vijayakumar, Thomas Niederkrotenthaler, Mala Jayaseelan, Ramya Kannan, Jane Pirkis, and Anthony Jorm. 2018. Assessing the quality of Media Reporting of Suicide News in India against World Health Organization Guidelines: A Content Analysis Study of Nine Major Newspapers in Tamil Nadu. *Australian & New Zealand Journal of Psychiatry* 52, 9: 856–863.

[14]   Cathy Babao. 2018. What Parents Must Know Now: "13 Reasons Why," or Why More Youngsters are
       Committing Suicide.

[15]   Albert Bandura, Edward Blanchard, and Ritter Brunhilde. 1969. Relative Efficacy of Desensitization and
       Modeling Approaches for Inducing Behavioral, Affective, and Attitudinal Changes. *Journal of
       Personality and Social Psychology* 13, 3: 173–199.

[16]   Susan Beaton, Peter Forster, and Myfanwy Maple. 2013. *Suicide and Language: Why We Shouldn't Use
       the "C" Word.*

[17]   India Bohanna and Xiangdong Wang. 2012. Media Guidelines for the Responsible Reporting of Suicide:
       A Review of Effectiveness. *Crisis: The Journal of Crisis Intervention and Suicide Prevention* 33, 4: 190–
       198.

[18]   Virginia Braun and Victoria Clarke. 2008. Using Thematic Analysis in Psychology. *Qualitative Research
       In Psychology* 3, 2: 77–101.

[19]   Sean Brotherson and April Anderson. 2016. Talking to Children About Suicide. *North Dakota State
       University.*

[20]   Zee Bureau. 2017. Consider Taking Steps to Firewall Games Like Blue Whale, SC Tells Govt. *Zee News.*

[21]   Pete Burnap, Walter Colombo, and Jonathan Scourfield. 2015. Machine Classification and Analysis of
       Suicide-Related Communication on Twitter.

[22]   Victoria Carmichael and Rob Whitley. 2019. Media Coverage of Robin Williams' Suicide in the United
       States: A Contributor to Contagion? *PloS One* 14, 5.

[23]   Kapil Chalil Madathil, Joy Rivera-Rodriguez, Joel Greenstein, and Anand Gramopadhye. 2015.
       Healthcare information on Youtube: A systematic review. *Health Informatics Journal* 21, 3: 173–194.

[24]   Man-pui Chan, Christopher Jones, Kathleen Jamieson, and Dolores Albarracín. Debunking: A Meta-
       Analysis of the Psychological Efficacy of Messages Countering Misinformation. *Psychological
       Science* 28, 11: 1531–1546.

[25]   Stevie Chancellor, Zhiyuan Lin, and Munmun De Choudhury. 2016. "This Post Will Just Get Taken
       Down": Characterizing Removed Pro-Eating Disorder Social Media Content. *Computer Supported
       Cooperative Work.*

[26]   Stevie Chancellor, Zhiyuan Lin, Erica Goodman, Stephanie Zerwas, and Munmun De Choudhury. 2016.
       Quantifying and Predicting Mental Illness Severity in Online Pro-Eating Disorder Communities.
       Computer Supported Cooperative Work.

[27]   Stevie Chancellor, Jessica Pater, Trustin Clear, Eric Gilbert, and Munmun De Choudhury. 2016.
       #thyghgapp: Instagram Content Moderation and Lexical Variation in Pro-Eating Disorder
       Communities. *Computer Supported Cooperative Work.*

[28]   Prabha Chandra, Padmavathy Doraiswamy, Anuroopa Padmanabh, and Mariamma Philip. 2014. Do
       Newspaper Reports of Suicides Comply with Standard Suicide Reporting Guidelines? A Study from
       Bangalore, India. *International Journal of Social Psychiatry* 60, 7: 687–694.

[29]   Allyson Chiu. 2018. The "Momo Challenge": A Sinister Threat to Young People or an Urban Myth?
       Washington Post Blogs.

[30]   Munmun De Choudhury. 2015. Anorexia on Tumblr: A Characterization Study. *DH '15 Proceedings of
       the 5th International Conference on Digital Health*: 43–50.

[31]   Munmun De Choudhury and Sushovan De. 2014. Mental Health Discourse on Reddit: Self-Disclosure,
       Social Support, and Anonymity. *Proceedings of the Eighth International AAAI Conference on Weblogs
       and Social Media.*

[32]   Munmun De Choudhury, Emre Kiciman, Mark Dredze, Glen Coppersmith, and Mrinal Kumar. 2016.
       Discovering Shifts to Suicidal Ideation from Mental Health Content in Social Media. *Proceedings of
       the SIGCHI Conference on Human Factors in Computing Systems*: 2098–2110.

[33]   Unmana Dutta. 2017. Online Trauma: No Child's Game. *Alive.*

[34]   Hend El-behary. 2018. 12-year-old Egyptian Schoolboy Commits Suicide to Fulfill "Blue Whale"
       Challenge. *Egypt Independent.*

[35]   Elmar Etzersdorfer and Gernot Sonneck. 1998. Preventing Suicide by Influencing Mass-Media
       Reporting: The Viennese Experience 1980-1996. *Archives of Suicide Research* 4, 1: 67–74.

[36] Emilio Ferrara, Onur Varol, Clayton Davis, Filippo Menczer, and Alessandro Flammini. 2016. The Rise of Social Bots. *Communications of the ACM* 59, 7: 96–104.

[37] Manas Gaur, Amanuel Alambo, Joy Sain, Ugur Kursuncu, K Thirunarayan, Ramakanth Kavuluru, Amit Sheth, Randy Welton, and Jyotishman Pathak. 2019. Knowledge-Aware Assessment of Severity of Suicide Risk for Early Intervention.

[38] Madelyn Gould. 1990. Suicide Clusters and Media Exposure. In *Suicide over the life cycle: Risk factors, assessment, and treatment of suicidal patients*. American Psychological Association, Arlington, VA, 517–532.

[39] Madelyn Gould, Patrick Jamieson, and Daniel Romer. 2003. Media Contagion and Suicide Among the Young. *American Behavioral Scientist* 46, 9: 1269–1284.

[40] Aditi Gupta. 2015. Designing and evaluating techniques to mitigate misinformation spread and microblogging web services. Indraprastha Institute of Information Technology Delhi.

[41] Camilla Haw, Keith Hawton, Claire Niedzwiedz, and Steve Platt. Suicide Clusters: A Review of Risk Factors and Mechanisms. *Suicide and Life-Threatening Behavior* 43, 1: 97–108.

[42] Charlotte Hilton. 2017. Unveiling self-harm behaviour: what can social media site Twitter tell us about self-harm? A qualitative exploration. *Journal of Clinical Nursing* 26, 11–12: 1690–1704. https://doi.org/10.1111/jocn.13575

[43] Nikhil Jain and Sunil Kumar. 2016. Is Suicide Reporting in Indian Newspapers Responsible? A Study from Rajasthan. *Asian Journal of Psychiatry* 24: 135–138.

[44] Haiyan Jia, Pamela Wisniewski, Heng Xu, Mary Rosson, and John Carroll. 2015. Risk-Taking as a Learning Process for Shaping Teen's Online Information Privacy Behaviors. *Computer Supported Cooperative Work*.

[45] Andreas Kaplan and Michael Haenlein. 2010. Users of the World, Unite! The Challenges and Opportunities of Social Media. *Business Horizons* 53, 1: 59–68.

[46] Amro Khasawneh, Kapil Chalil Madathil, Emma Dixon, Pamela Wisniewski, Heidi Zinzow, and Rebecca Roth. 2019. An Investigation on the Portrayal of Blue Whale Challenge on YouTube and Twitter. In (1), 887–888.

[47] Abhinav Khattar, Karan Dabas, Kshitij Gupta, Shaan Chopra, and Ponnurangam Kumaraguru. 2018. White or Blue, the Whale Gets its Vengeance: A Social Media Analysis of The Blue Whale Challenge. Retrieved from https://arxiv.org/abs/1801.05588

[48] Bonnie Kilmes-Dougan, Chih Yuan, and Alaa Houri. 2009. Suicide Prevention with Adolescents Considering Potential Benefits and Untoward Effects of Public Service Announcements. *Crisis: The Journal of Crisis Intervention and Suicide Prevention* 30: 128–135.

[49] Sabine Kunrath, Jens Baumert, and Karl-Heinz Ladwig. 2010. Increasing Railway Suicide Acts After Media Coverage of a Fatal Railway Accident? An Ecological Study of 747 Suicidal Acts. *Journal of Epidemiology and Community Health* 65, 9: 825–828.

[50] Daria Kuss and Mark Griffiths. 2012. Internet Gaming Addiction: A Systematic Review of Empirical Research. *International Journal of Mental Health and Addiction* 10, 2: 279–296.

[51] Simone Lanette and Melissa Mazmanian. The Smartphone "Addiction" Narrative is Compelling, but Largely Unfounded.

[52] Stephan Lewandowsky, Ullrich Ecker, Colleen Seifert, Norbert Schwarz, and John Cook. 2012. Misinformation and Its correction: Continued influence and successful debiasing. *Psychological Science in Public Interest: American Psychological Society* 13, 3: 106–131.

[53] Lupariello. Self-Harm Risk Among Adolescents and the Phenomenon of the "Blue Whale Challenge": Case Series and Review of the Literature.

[54] Francesco Lupariello, Serena Curti, Elena Coppo, Sara Racalbuto, and Giancarlo Di Vella. 2019. Self-harm Risk Among Adolescents and the Phenomenon of the "Blue Whale Challenge": Case Series and Review of the Literature. *Journal of Forensic Science* 64: 638–642.

[55] Myfanwy Maple, Julie Cerel, Rebecca Sanford, Tania Pearce, and Jack Jordan. 2017. Is Exposure to Suicide Beyond Kin Associated with Risk for Suicidal Behavior? A Systematic Review of the Evidence. *Suicide and Life-Threatening Behavior* 47, 4: 461–474.

[56] Megan Moreno, Aubrey Gower, Heather Brittain, and Tracy Vaillancourt. 2019. Applying Natural Language Processing to Evaluate News Media Coverage of Bullying and Cyberbullying. *Prevention Science* 20.

[57] Richa Mukhra, Neha Baryah, Kewal Krishan, and Tanuj Kanchan. 2017. 'Blue Whale Challenge': A Game or Crime? *Science and Engineering Ethics*. https://doi.org/10.1007/s11948-017-0004-2

[58] Richa Mukhra, Neha Baryah, Krishan Krishan, and Tanuj Kanchan. 2007. "Blue Whale Challenge": A Game or a Crime? *Science and Engineering Ethics*.

[59] National Action Alliance for Suicide Prevention. *Tips for Messaging Safely*. Retrieved from http://suicidepreventionmessaging.org/safety/tips-messaging-safely

[60] National Institute of Mental Health. 2015. NIMH » Suicide. Retrieved November 16, 2017 from https://www.nimh.nih.gov/health/statistics/suicide/index.shtml

[61] Thomas Niederkrotenthaler, Daniel Reidenberg, Benedikt Till, and Madelyn Gould. 2014. Increasing Help-Seeking and Referrals for Individuals at Risk for Suicide by Decreasing Stigma: The Role of Mass Media. *American Journal of Preventitive Medicine* 47, 3: 235–243.

[62] Thomas Niederkrotenthaler and Gernot Sonneck. 2007. Assessing the Impact of Media Guidelines for Reporting on Suicides in Austria: Interrupted Time Series Analysis. *Australian and New Zealand Journal of Psychiatry* 41, 5: 419–428.

[63] Thomas Niederkrotenthaler, Martin Voraek, Arno Herberth, Benedikt Till, Markus Strauss, Elmar Etzersdorfer, and Gernot Sonneck. 2010. Role of Media Reports in Completed and Prevented Suicide: Werther v. Papageno Effects. *The British Journal of Psychiatry* 197, 3: 234–243.

[64] Alicia Nobles, Jeffrey Glenn, Kamran Kowsari, Bethany Teachman, and Laura Barnes. 2018. Identification of Imminent Suicide Risk Among Young Adults using Text Messages.

[65] Haim Omer and Avshalom Elitzur. 2001. What Would You Say to the Person on the Roof? A Suicide Prevention Text. *Suicide and Life-Threatening Behavior* 31.

[66] Amy Orben and Andrew Przybylski. 2019. The association between adolescent well-being and digital technology use. *Nature: Human Behavior* 3: 173–182.

[67] Jessica Pater, Brooke Farrington, Alycia Brown, Lauren Reining, Tammy Toscos, and Elizabeth Mynatt. Exploring Indicators of Digital Self-Harm with Eating Disorder Patients: A Case Study. *Computer Supported Cooperative Work* 3.

[68] Jessica Pater, Oliver Haimson, Nazanin Andalibi, and Elizabeth Mynatt. 2016. "Hunger Hurts but Starving Works:" Characterizing the Presentation of Eating Disorders Online! *Computer Supported Cooperative Work*.

[69] Jessica Pater and Elizabeth Mynatt. 2017. Defining Digital Self-Harm. *Computer Supported Cooperative Work*.

[70] David Phillips. 1979. Suicide, Motor Vehicle Fatalities, and the Mass Media: Evidence Toward a Theory of Suggestion. *American Journal of Sociology* 84, 5: 1150–1174.

[71] David Phillips and Katherine Lesyna. 1995. Suicide and the Media: Research and Policy Implications. In *Preventative Strategies on Suicide*. 231–261.

[72] David Phillips, Katherine Lesyna, and Daniel Paight. 1992. Suicide and the Media. In *Assessment and prediction of suicide*. Guilford Press, New York, NY, US, 499–519.

[73] Wayne Rhee, Michael Merbaum, Michael Strube, and Susan Self. 2005. Efficacy of Brief Telephone Psychotherapy with Callers to a Suicide Hotline. *Suicide and Life-Threatening Behavior* 35, 3: 317–328.

[74] Jingyuan Shi and Hye Kim. 2019. Integrating Risk Perception Attitude Framework and the Theory of Planned Behavior to Predict Mental Health Promotion Behaviors Among Young Adults. *Health Communication*: 1–10.

[75] Craig Silverman. 2015. Lies, damn lies, and viral content. How news websites spread (and debunk) online rumors, unverified claims, and misinformation. *Tow Center for Digital Journalism* 168.

[76] David Silverman. 2015. *Interpreting Qualitative Data*. SAGE Publications.

[77] Sandra Song. 2019. Internet Hoax Hyesteria: What's the Truth? *Paper Magazine*.

[78] Steven Stack. 2000. Media Impacts on Suicide: A Quantitative Review of 293 Findings. *Social Science Quarterly* 81, 4: 957–971.

[79] Kevin Stanford. 2016. Social Media Contests: A Timeline of Viral Challenges. *Growing Social Media.* Retrieved October 20, 2017 from http://growingsocialmedia.com/social-media-contests-a-timeline-of-viral-challenges/

[80] Kate Starbird, Dharma Dailey, Owla Mohamed, Gina Lee, and Emma Spiro. 2018. Engage Early, Correct More: How Journalists Participate in False Rumors Online during Crisis Events. *Computer Supported Cooperative Work.*

[81] Roma Subramanian. 2018. Frames of Mental Illness in an Indian Daily Newspaper. *Health Communication*: 1–10.

[82] Suicide Prevention Resource Center. 2006. *Safe and Effective Messaging for Suicide Prevention.*

[83] Steven Sumner, Stacey Galik, Jennifer Mathieu, Megan Ward, Thomas Kiley, Brad Bartholow, Alison Dingwall, and Peter Mork. 2019. Temporal and Geographic Patterns of Social Media Posts About an Emerging Suicide Game. *Journal of Adolescent Health*, 65: 94–100.

[84] Sonja Swanson and Ian Colman. 2013. Association Between Exposure to Suicide and Suicidality Outcomes in Youth. *Canadian Medical Association Journal* 185, 10: 870–877.

[85] Jean M. Twenge, Thomas E. Joiner, Megan L. Rogers, and Gabrielle N. Martin. 2017. Increases in Depressive Symptoms, Suicide-Related Outcomes, and Suicide Rates Among U.S. Adolescents After 2010 and Links to Increased New Media Screen Time. *Clinical Psychological Science*: 2167702617723376. https://doi.org/10.1177/2167702617723376

[86] Alicia VanOrman and Beth Jarosz. 2016. Suicide Replaces Homicide as Second-Leading Cause of Death Among U.S. Teenagers. *Population Reference Bureau.* Retrieved November 16, 2017 from http://www.prb.org/Publications/Articles/2016/suicide-replaces-homicide-second-leading-cause-death-among-us-teens.aspx

[87] Yilin Wang, Jiliang Tang, Jundong Li, Baoxin Li, Yali Wan, Clayton Mellina, Neil O'Hare, and Yi Chang. Understanding and Discovering Deliberate Self-harm Content in Social Media. *Computer Supported Cooperative Work.*

[88] Pamela Wisniewski. 2018. The Privacy Paradox of Adolescent Online Safety: A Matter of Risk Prevention or Risk Resilience? *Security & Privacy* 16: 86–90.

[89] Pamela Wisniewski, Arup Ghosh, Heng Xu, Mary Rosson, and John Carroll. 2017. Parental Control vs. Teen Self-Regulation: Is There a Middle Ground for Mobile Online Safety? *Computer Supported Cooperative Work*: 51–69.

[90] Pamela Wisniewski, Haiyan Jia, Na Wang, Saijing Zheng, Henry Xu, Mary Rosson, and John Carroll. Resilience Mitigates the Negative Effects of Adolescent Internet Addiction and Online Risk Exposure. *ACM Conference on Human Factors in Computing Systems.*

[91] Pamela Wisniewski, Henry Xu, Mary Rosson, Daniel Perkins, and John Carroll. "Dear Diary: Teens Reflect on Their Weekly Online Risk Experience.

[92] World Health Organization. 2017. *Preventing Suicide: A Resource for Media Professionals, Update 2017.*

[93] Albert Yang, Shih Tsai, Cheng Yang, Ben-Chang Shia, Jong Fuh, Shuu Wang, Chung Peng, and Norden Huang. Suicide and Media Reporting: A Longitudinal and Spatial Analysis. *Social Psychiatry and Psychiatric Epidemiology* 48, 3: 427–435.

[94] 2015. *Recommendations for Reporting on Suicide.* Retrieved from reportingonsuicide.org

[95] 2016. *Suicide Mortality Rate.* World Health Organization. Retrieved from http://apps.who.int/gho/data/node.sdg.3-4-viz-2?lang=en

[96] 2017. The Blue Whale Challenge: Fact or Fiction? *Global English.*

[97] 2017. ALERT: Parents of Dead Teen Find Clues to Her Suicide On Bedroom Wall - It's a Deadly Game, and It's Spreading. *Right Wing News.*

[98] 2017. Class 10 Panchkula Student Found Hanging, Family Blames Blue Whale. *Hindustan Times.*

[99] 2017. Blue Whale Challenge to Take Lives. *Afternoon Voice.*

[100] 2017. Odisha Girl "Playing" Blue Whale Challenge Attempts Suicide. *Odisha Times.*

[101] 2017. Texas Family Says Teen Killed Himself in Macabre "Blue Whale" Online Challenge that's Alarming Schools. *Washington Post Blogs.*

[102] 2017. How to Protect Kids from Blue Whale? *Global Times China.*

[103] 2017. Govt Tells Tech Giants to Remove "Blue Whale" Game Links. *Oman Tribune.*
[104]2017. Dangerous Games: The Blue Whale Challenge, the Black Hole of the Cyber World. *Alive.*
[105]2017. Blue Whale Challenge: Parents Told to Keep an Eye on Children's Online Activity. *International Business Times India.*
[106]2017. Another Suicide Suspected due to Blue Whale Game in Kerala. *Business Line.*
[107]2017. Parents, Suburb Kids Jack up Blue Whale Hits Online. *Times of India.*
[108]2017. Blue Whale: Jodhpur Girl Attempts Suicide Again, Rushed to Hospital. *DNA India.*
[109]2018. *Lexis Nexis.* Retrieved from https://www.lexisnexis.com/en-us/gateway.page
[110]Russia: Man Hehind Blue Whale Suicide "Game" is Jailed.
[111]*Fake News & Fact Checking.* University of Central Florida Libraries. Retrieved from https://guides.ucf.edu/fakenews